

WHAT IS CLAIMED IS:

1. In a system comprised of a plurality of storage elements, a method for maintaining objects in the storage elements comprising the steps of:

5 maintaining information regarding which storage elements are storing particular objects in a consistency coordinator which communicates with the storage elements;

responding to a request to update an object by using maintained information to determine which of the storage
10 elements may store a copy of the object;

instructing the storage elements, which the consistency coordinator suspects store a copy of the object, to invalidate their copy of the object; and

performing an update of the object after each storage
15 element that includes the copy of the object indicates that the storage element has invalidated the copy of the object or the storage element is determined to be unresponsive.

2. The method as recited in claim 1, wherein the
20 step of maintaining information includes maintaining information regarding which storage elements are storing particular objects in the consistency coordinator.

3. The method as recited in claim 1, wherein the consistency coordinator includes multiple nodes and each node of the consistency coordinator stores information for a different set of objects.

5

4. The method as recited in claim 1, wherein the storage elements include at least one cache.

10

5. The method as recited in claim 1, wherein the storage elements are included in a distributed system.

6. The method as recited in claim 1, further comprising the step of obtaining a lock on the object to be updated before performing the update.

15

7. The method as recited in claim 1, further comprising the step of sending heart beat messages to obtain availability information about objects from the maintained information to a storage element and from a storage element to the maintained information.

20

8. The method as recited in claim 7, further comprising the step of declaring an entity down in response

to failing to receive a heart beat.

9. The method as recited in claim 7, wherein the entity declares itself down in response to failing to
5 receive a heart beat.

10. A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for maintaining strong
10 data consistency, the method steps comprising:

maintaining information regarding which storage elements are storing particular objects in a consistency coordinator which communicates with the storage elements;

responding to a request to update an object by using
15 maintained information to determine which of the storage elements may store a copy of the object;

instructing the storage elements, which the consistency coordinator suspects store a copy of the object, to invalidate their copy of the object; and

20 performing an update of the object after each storage element that includes the copy of the object indicates that the storage element has invalidated the copy of the object or the storage element is determined to be unresponsive.

11. In a system comprised of a plurality of storage elements, a method for maintaining stored objects comprising the steps of:

5 maintaining a consistency coordinator which communicates with the storage elements and stores information regarding which storage elements are storing which objects;

10 in response to receiving a request to update an object, using information from the consistency coordinator to determine a set of storage elements which may store a copy of the object;

 instructing each storage element in the set to invalidate a copy of the object; and

15 performing the update after each storage element in the set indicates that the storage element has invalidated a copy of the object or the storage element is determined to be unresponsive.

20 12. The method as recited in claim 11, wherein the consistency coordinator includes multiple nodes and further comprising the step of at each node of the consistency coordinator, storing information about which storage

elements are storing which objects for a different set of objects.

13. The method as recited in claim 11, further
5 comprising obtaining a lock from the consistency coordinator by an entity attempting to update an object before performing the update.

14. The method as recited in claim 11, further
10 comprising the step of sending, from the consistency coordinator to a storage element or from a storage element to the consistency coordinator, heart beat messages to obtain availability information.

15. The method as recited in claim 14, further
15 comprising an entity expecting a heart beat, declaring itself down in response to failing to receive a heartbeat.

16. The method as recited in claim 11, wherein the
20 storage elements include at least one cache.

17. A program storage device readable by machine, tangibly embodying a program of instructions executable by

the machine to perform method steps for maintaining strong data consistency, the method steps comprising:

maintaining a consistency coordinator which communicates with the storage elements and stores information regarding which storage elements are storing which objects;

in response to receiving a request to update an object, using information from the consistency coordinator to determine a set of storage elements which may store a copy of the object;

instructing each storage element in the set to invalidate a copy of the object; and

performing the update after each storage element in the set indicates that the storage element has invalidated a copy of the object or the storage element is determined to be unresponsive.

18. A system for maintaining strong data consistency comprising:

a plurality of storage elements; and

a consistency coordinator, which communicates with the plurality of storage elements and maintains information about which objects are stored in the plurality of storage

elements,

the consistency coordinator providing selective communication to storage elements which include an object to be updated such that for a given object update the consistency coordinator communicates with only those storage elements which include the object to be updated.

19. The system as recited in claim 18, further comprising a writer, which updates the object to be updated.

20. The system as recited in claim 19, wherein the writer resides on a same node as a storage element.

21. The system as recited in claim 19, wherein the writer writes an updated object to storage elements after the plurality of storage elements which are to receive the update have invalidated a current copy of the object.

22. The system as recited in claim 19, wherein the writer writes an updated object to storage elements after the plurality of storage elements which are to receive the update are determined to be unresponsive.

23. The system as recited in claim 18, further comprising at least one content provider.

5 24. The system as recited in claim 23, wherein the content provider resides on a same node as a storage element.

10 25. The system as recited in claim 18, further comprising heart beat messages, which may be transmitted between the consistency coordinator and the storage elements to obtain availability information from the consistency coordinator to a storage element or from a storage element to the consistency coordinator.

15

26. The system as recited in claim 18, wherein the storage elements include at least one cache.

20